DTU Library

Condition monitoring with ordinary wind turbine SCADA data – A neuro-fuzzy approach

Schlechtingen, Meik; Santos, Ilmar

Publication date:
2012

Citation (APA):
Schlechtingen, M., & Santos, I. (2012). *Condition monitoring with ordinary wind turbine SCADA data – A neuro-fuzzy approach*. Paper presented at Offshore 2011, Amsterdam, Netherlands.

Condition monitoring with ordinary wind turbine SCADA data – A neuro-fuzzy approach

Meik Schlechtingen
Department of Technical Operation Wind Offshore,
EnBW Erneuerbare Energien GmbH, Admiralitätstr.4, 20459 Hamburg, Germany
Tel.: +49 40533268134, Email: m.schlechtingen@enbw.com

Ilmar Ferreira Santos
Department of Mechanical Engineering, Section of Solid Mechanics,
Technical University of Denmark, Nils Koppels Allé Building 404, 2800 Kgs. Lyngby, Denmark
+45 45931577, Email: ifs@mek.dtu.dk

Abstract
This paper presents the latest research results of a project that focuses on normal behavior models for condition monitoring of wind turbine components and the turbine itself, via ordinary Supervisory Control And Data Acquisition (SCADA) data. In this machine learning approach Adaptive Neuro-Fuzzy Interference System (ANFIS) models are employed to learn the normal behavior in a training phase, where the component condition can be considered healthy. In the application phase the trained models are applied to predict the target signals, e.g. temperatures, pressures, currents, etc.
The behavior of the prediction error is used as an indicator, whether the monitored signal behaves normal or abnormal, with respect to the learned behavior.
For classification of the component condition Fuzzy Interference System (FIS) structures are used. Based on rules that are established with the prediction error behavior during previously experienced faults and generic rules the FIS structures output the component condition (green, yellow and red). Furthermore a first diagnosis of the root cause is given. First results of a field test show the applicability of the approach to closely monitor a large variety of signals, identify anomalies and classify the component condition correctly. Within this paper two examples of real measured faults are given that show the capabilities of the proposed method. The method can be applied both to existing and new built turbines without additional hardware installation.

Introduction
With wind turbines built in remote locations, reliable turbine and component condition statements become increasingly important. Especially offshore a break down maintenance strategy can result in high costs due to turbine downtimes and non-availability of large components and specialized vessels like crane ships. Condition Monitoring Systems (CMS) aim to provide operators with information regarding the “health” of their machines, which in turn, can help them improve operational efficiency by allowing more informed decisions regarding maintenance [1-7]. There is a large variety of CMS developed during the past years that focus on vibration diagnosis of wind turbines. On the other hand the ordinary SCADA data have long been used only for backward fault analysis after fault occurrence. The available CMS mostly require high level knowledge about the system to be monitored. This knowledge is difficult to access and does often not exist. Physical models of the system, to monitor its condition and predict failures can thus seldom be built with high accuracy due to its complex interaction among several dynamical subsystems.
However, there is a large amount of operational data available in data bases, which can be used to give an indication on the turbine condition. In [1,6,8-10] it is shown that by use of artificial intelligence approaches information can be extracted that allows diagnosis of incipient fault days, weeks and in some cases even months in advance. In the aforementioned works artificial neural networks are employed to develop Normal Behaviour Models (NBM) at a state, where the component or turbine can be considered healthy. Usually this state is assumed to be at the beginning of the turbine or component lifetime. The learned normal behaviour can then be used in the application phase to examine deviations from the normal behavior (in terms of the prediction error). The great advantage of this approach is that the prediction error is widely decoupled from the typical fluctuations of the SCADA data caused by the different turbine

operational modes.
In this paper it is described a CMS for wind turbines that takes advantage of Adaptive Neuro Fuzzy Interference Systems (ANFIS) to develop the NBM. ANFIS models can learn nonlinear signal relations by setting up a set of fuzzy rules and tuning the membership function parameters in a training phase. The application of ANFIS models for CMS purposes presents a novelty. The approach is applied to a broad variety of SCADA signals coming from an onshore wind farm consisting of 18 turbines of the 2MW class.
When monitoring a broad variety of SCADA signals of a turbine, it can be realized that certain patterns and interactions establish within the prediction errors of the different models. For example in case of unexpected low turbine power output, not only the power output prediction error will indicate this anomaly, but also the generator current. These patterns can be used to diagnose the root cause of a problem. The rules that apply in diagnosing the patterns can be built with fuzzy logic. Fuzzy logic allows implementation of human intuition in terms of rules in a straight forward manner. Furthermore it allows solutions between the hard boundaries true or false when applying a rule. This characteristic let the fuzzy logic structure even output a solution if a rule only partially applies. In the CMS fuzzy logic is used to link the prediction error of different signals to arrive at a condition statement.
In the following the general CMS concept is described and two application examples as well as a list of processed signals are given.

Description of the general CMS concept

The CMS aims to detect trends and patterns in SCADA data in order to identify incipient faults at an early stage. This stage should leave the operator enough time to adapt the maintenance schedule or take further measures to prevent unexpected system downtime. The CMS processes and monitors 10 minute average values such as temperatures, currents or the power output. This is done by taking advantage of the dependencies between the signals and set up NBM. A simple example showing the working principle of NBM development is the relation between the generator bearing and the generator stator temperature. The bearing temperature is mostly determined by external influences and can thus be expressed by the stator temperature, the shaft speed, the power output and the nacelle temperature. Taking these four signals, the generator bearing temperature can be reconstructed (see Figure 1) by training a model the relation between the input and the target signal. However, the dependency between the signals is often nonlinear and thus well suited to be learned by ANFIS models.

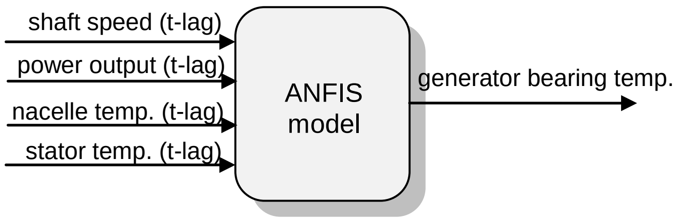


Figure 1: Model working principle

The reconstructed signal or expectation can be compared with the real measured signal and the prediction error calculated and monitored. It is worth noting that the CMS uses offline data and hence real time analysis of the SCADA data is not covered in this research. The ability to analyze faults in real time is limited by the fact that there exists a signal response delay that can be up to 60 min. An example for such a delay is the gearbox oil temperature; if the power output of the turbine increases, the gearbox oil temperature follows this increase but with a delay of up to 60 min., due to the large heat capacity of the component. Furthermore, real time SCADA data are currently seldom available to operators due to both technical and historical reasons.
Taking into account the restrictions on data availability, the monitoring methods developed in this research use 10 min. average values that are usually available. Often these data are submitted sequentially with some time delay, e.g. 4 h, which limits the possibility of detecting very fast propagating faults. In earlier work of the authors [10] it was found that averaging the prediction error to one day average values not only decreases the variance of the prediction error, it also increases the fault sensitivity. It means that by averaging to one day averages faults are easier and earlier to identify. For this reason the CMS outputs a single prediction error

per signal per day that is used for further analysis.

The general system architecture of the CMS developed is shown in Figure 2.

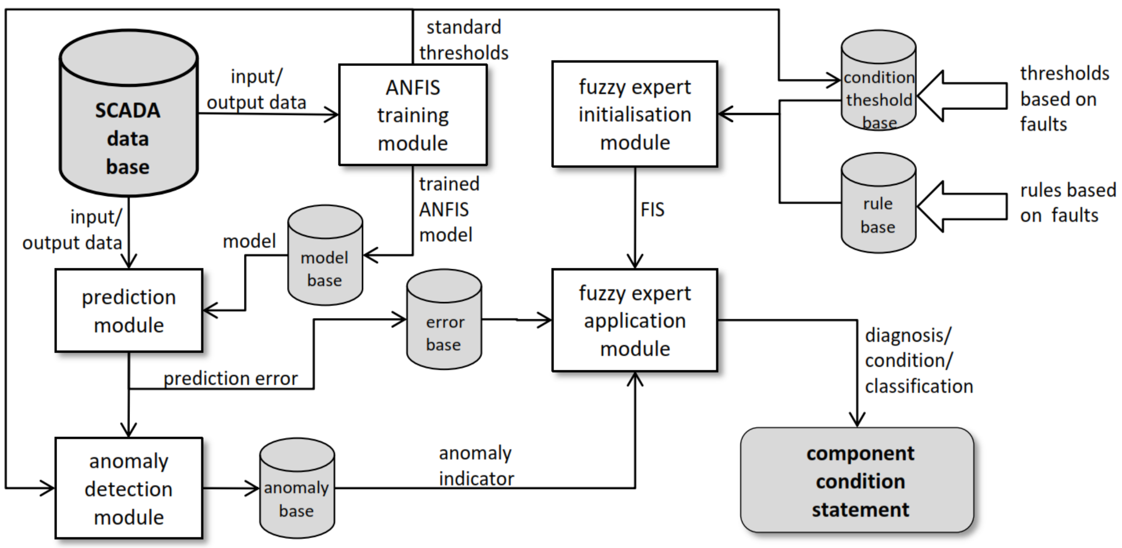


Figure 2: CMS overview

In the following, the function of the different CMS modules (see Figure 2) is briefly described. ***Training module***: Within the training module the NBM is trained if a model is not yet available or new training is required. The latter is true if a component is replaced and the signal relations change as a consequence. Before training the model, the data are preprocessed according to the methodology proposed in [10] which includes: 1) a validity check, 2) data range check, 3) missing data processing and 4) lag removal.

In the training module different training levels are implemented to allow early monitoring. A first model training is performed after one month of operational data collection. Further trainings are performed when three, six and nine months of data are available. The output of the training module is the trained ANFIS model and the standard thresholds marking the normal range of the prediction error.

Prediction module: The prediction module is active once a trained model of the processed signal is available in the model base. The developed NBM is applied and the prediction error calculated and stored.

Anomaly detection module: In this module the anomalies in the prediction errors are identified. This is done on the basis of the determined normal-behavior thresholds by the training module or expert defined values. The output is an anomaly matrix containing information about the frequency and date of occurrence, as well as the duration of the current anomaly in days.

Fuzzy expert initialization module: Here the FIS structure used for diagnosis of the anomalies and component condition statements are initialized. Each FIS structure has the following outputs:

- diagnosis (information about the abnormal behaving signal)
- condition (classification in green, yellow and red color code) (grey if no diagnosis is possible)
- diagnosis certainty
- potential root cause

Fuzzy expert application module: Within this module the initialized FIS structure is evaluated, given the prediction errors and the information about present anomalies. The output is stored in text format and is visualized to give the analyst a comprehensive summery of the turbine condition.

List of monitored SCADA signals
The 18 turbines in the field test supply the following signals that are used to monitor the turbine and components condition (the location of the sensors is visualized in Figure 3).

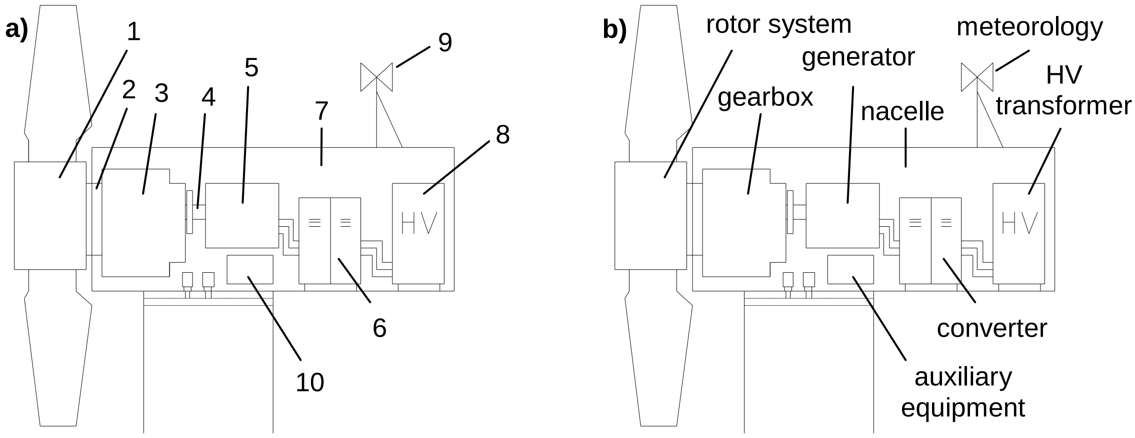


Figure 3: Wind turbine schematic: a) Sensor positions; b) Components or subsystems in the considered wind turbines

- Spinner temperature (pos.1)
- Hub controller temperature (pos.1)
- Pitch angle (pos.1)
- Hydraulic oil temperature (pos.10)
- Rotor speed (pos.2)
- Gear bearing temperature (HSS) (pos.3)
- Gear oil temperature (pos.3)
- Generator speed (pos.4)
- Generator bearing temperature1 (pos.5)
- Generator bearing temperature 2 (pos.5)
- Generator slip ring temperature (pos.5)
- Generator phase1 temperature (pos.5)
- Generator phase2 temperature (pos.5)
- Generator phase3 temperature (pos.5)
- Generator current phase1 (pos.6)
- Generator current phase2 (pos.6)
- Generator current phase3 (pos.6)
- Nacelle temperature (pos.7)
- Power output (pos. 6)
- Reactive power (pos. 6)
- Grid inverter phase1 temperature (pos. 6)
- Grid rotor inverter phase1 temperature (pos.6)
- Grid rotor inverter phase2 temperature (pos.6)
- Grid rotor inverter phase3 temperature (pos.6)
- Converter cooling water temperature (pos.6)
- Converter choke coil temperature (pos.6)
- Converter controller temperature (pos.6)
- Top controller temperature (pos.6)
- Grid busbar temperature (pos.8)
- HV transformer phase1 temperature (pos.8)
- HV transformer phase2 temperature (pos.8)
- HV transformer phase3 temperature (pos.8)
- Wind speed (pos.9)

Example 1
In this example an anomaly in hydraulic oil temperature is highlighted by the CMS developed. In addition to this it is emphasized how information about the anomaly root cause can be implemented in the CMS in order to ease the analysis of future faults. Figure 4 shows a 2D waterfall plot of the normalized averaged percentage prediction errors of the hydraulic oil temperature and the related inputs of the NBM over time. In this plot the colors indicate the prediction error amplitude. White areas mark periods where no prediction is available, e.g. due to missing data or non-operational periods.

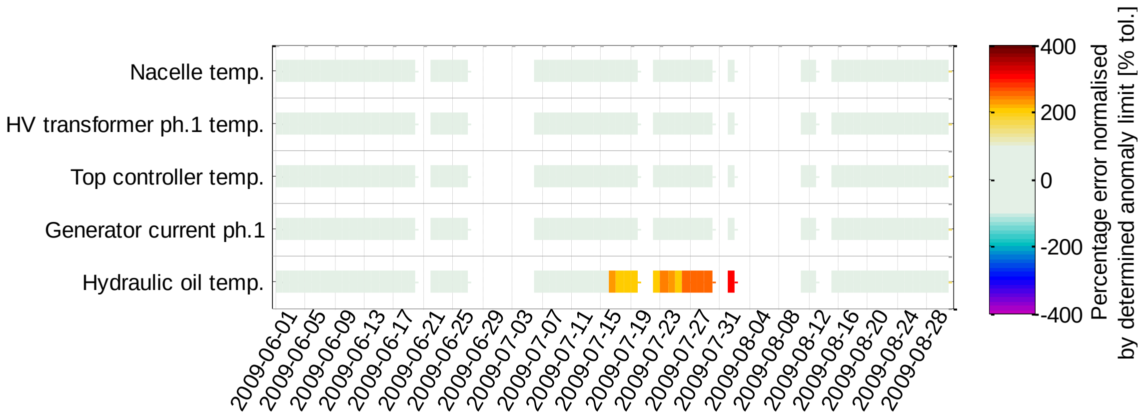
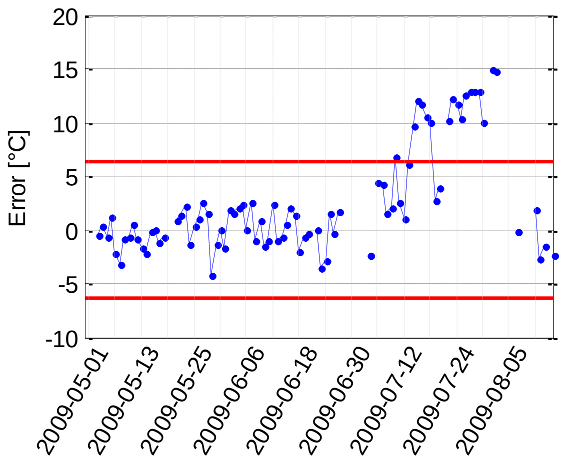


Figure 4: Normalized prediction errors of the hydraulic oil temperature and related inputs

It is clearly visible that beginning on 16.07.2011 an anomaly with rising amplitude occurs in the prediction error of the hydraulic oil temperature. Figure 5 shows the prediction error and the raw time series of the hydraulic oil temperature during fault occurrence over time to point out the amplitude height of the prediction error during fault occurrence.



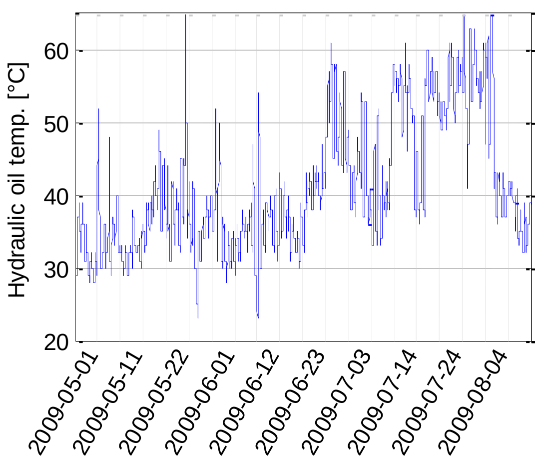


Figure 5: Hydraulic oil temperature prediction error and anomaly limits (left); hydraulic oil temperature raw time series (right)

The trend in the prediction error is clearly visible. Although the raw time series also indicates a trend, its visibility is dependent on the operational mode of the turbine.
The anomaly pattern of the normalized prediction error is analyzed by the FIS and a high hydraulic oil temperature is highlighted on 16.07.2009 by the CMS via the following generic rule:

If Hydraulic oil temp.==high & HV transformer ph.1 temp.==ok & Top controller temp.==ok & Generator current ph.1==ok & Nacelle temp.==ok then Diagnosis=Hydraulic oil temp. high; Condition=yellow; Pot. root cause=Ambiguous

This initial generic rule gives no information about the potential root cause and condition statement (yellow) is based on multipliers of the standard deviation instead of fixed temperature values. The fault can now be investigated and a rule established based on the diagnosed root cause. In this specific case, the root cause was a leakage in the rotary joint, causing the oil pump to run permanently. The entries in the service report were used to set up two specific rules containing the expert knowledge as follows.

1) If Hydraulic oil temp.==high & HV transformer ph.1 temp.==ok & Top controller temp.==ok & Generator current ph.1==ok & Nacelle temp.==ok then Diagnosis=Hydraulic oil temp. high; Condition=yellow; **Pot. root cause=Pump running permanently, leakage possible)**

2) If **Hydraulic oil temp.==very high** & HV transformer ph.1 temp.==ok & Top controller
temp. ==ok & Generator current ph.1==ok & Nacelle temp.==ok then Diagnosis=Hydraulic oil temp. high); **(Condition=red); (Pot. root cause=Pump running permanently, leakage possible)**

The terms high (yellow) and very high (red) of the hydraulic oil temp. prediction error are defined with 5°C and 14°C respectively. Reanalyzing the fault gives the results presented in Table 1.

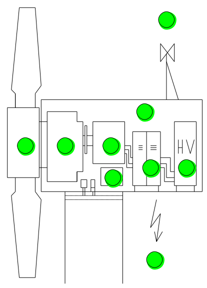
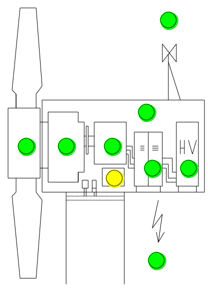
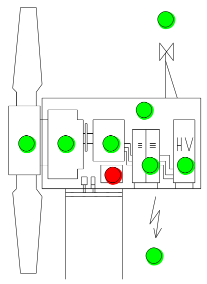
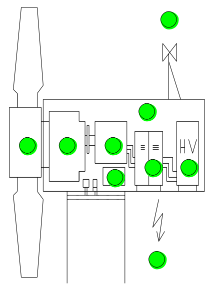
Date:	15.7.2009	16.07.2009	01.08.2009	11.08.2009
Condition:				
Diagnosis:	All components working as expected	Hydraulic oil temp. too high	Hydraulic oil temp. too high	All components working as expected
Pot. Root cause:	None	Pump running permanently leakage possible	Pump running permanently leakage possible	None
Certainty:	1.00	0.70	0.52	1.00

Table 1: Turbine condition evolution during hydraulic oil leakage occurrence

The expert knowledge was successfully implemented and the CMS is now able to identify and diagnose similar issues on other turbines of the fleet. It is worth noting, that the generic rules can be set up before fault occurrence based on general engineering knowledge about the system behavior in fault situations.

Example 2

The second example shows an increase in gearbox oil temperature due to dirty filters of the cooling system. Again a 2D waterfall plot is used to emphasize present anomaly patterns that establish during fault occurrence (see Figure 6). In this example the expert knowledge was implemented with a similar issue experienced earlier with another turbine.

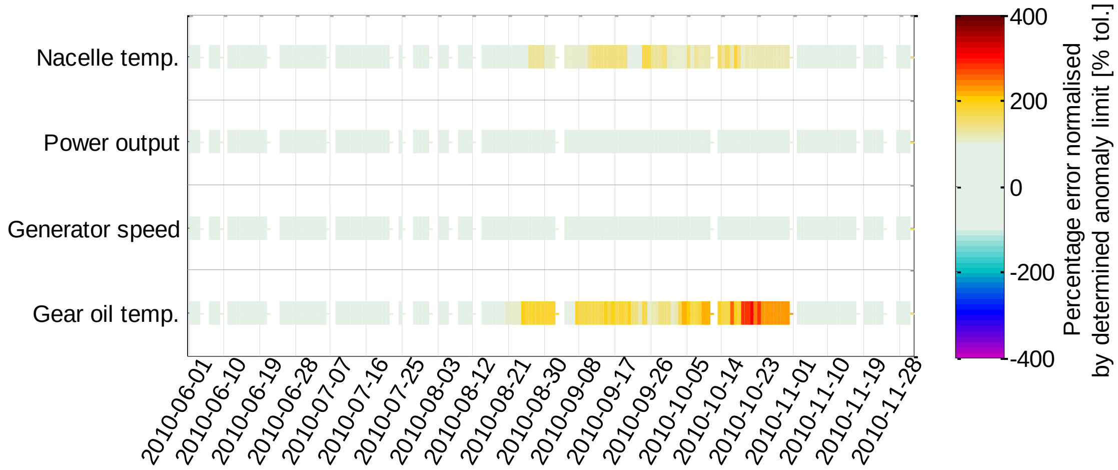


Figure 6: Normalized prediction errors of the gearbox oil temperature and related inputs

The visible high nacelle temperature is also a result of the dirty cooling system filters. Figure 7 shows the prediction error amplitudes and the raw time series of the gearbox oil temperature during fault occurrence.

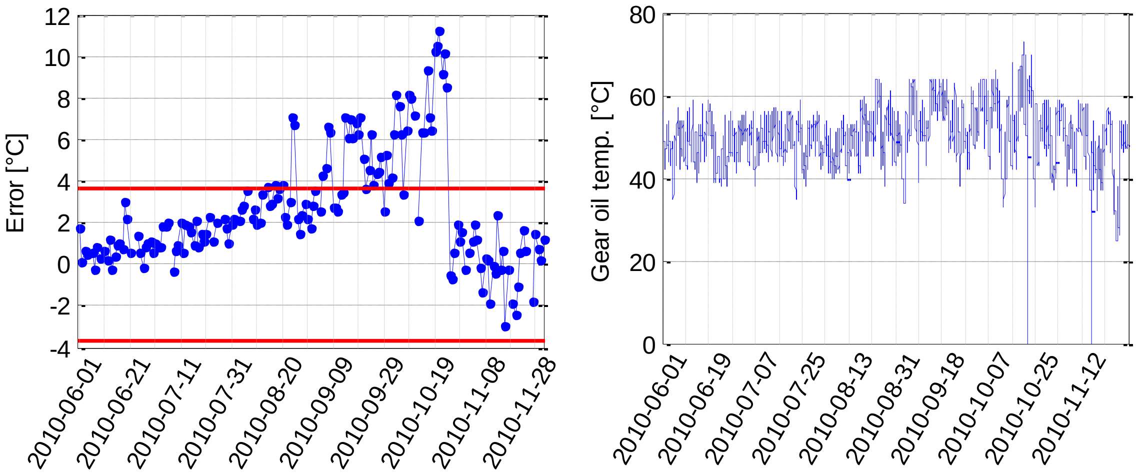


Figure 7: Gear oil temperature prediction error and anomaly limits (left); gear oil temperature raw time series (right)

Although the temperature deviation from normal behavior is up to 11°C the increase was not highlighted by the turbine controller, due to usually high thresholds set in the turbine controller. After filter replacement the temperature fell back to the normal operational range. The trend is also visible in the raw time series, but the fluctuations due to the different operational modes make it more difficult to identify the anomaly. Note that setting simple thresholds to the raw signal would make the identification of the increase dependent on the operational mode. A partially loaded turbine delays the fault discovery, which shows the effectiveness of the proposed method.

The evolution of the gearbox condition during filter degradation is visualized in Table 2.

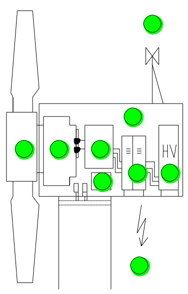
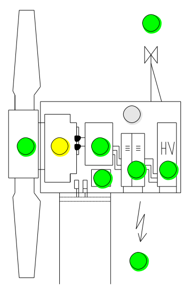
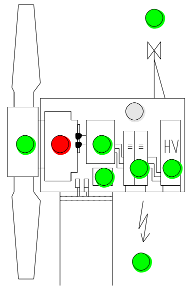
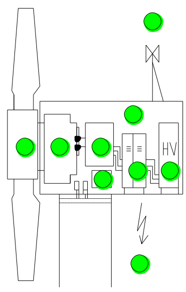
Date:	19.8.2010	20.08.2010	19.10.2010	20.10.2009
Condition:				
Diagnosis:	Component working as expected	Gearbox oil temp. too high	Gearbox oil temp. too high	All components working as expected
Pot. Root cause:	None	Cooling insufficient	Cooling insufficient	None
Certainty:	1.00	0.42	0.82	1.00

Table 2: Turbine condition evolution during gearbox oil temperature increase due to dirty filters

Conclusion

The research work underlines that SCADA data contain useful information about the component conditions. The developed CMS is capable of detecting incipient faults at an early stage as shown by the two examples in this paper. It works without any input from manufacturers and is able to detect anomalies even if they occur the first time. When an anomaly is detected the first time generic rules apply and highlight the issue. Fault analysis can take place and the gained expert knowledge can be implemented in terms of fuzzy rules in a straight forward manner. The research further emphasized that the specific rules implemented are valid for all turbines of the same type and thus allows diagnosis of similar faults on other turbines.

Problems with the approach occur, when no information about performed services is on hand and root cause analysis is impossible. In this case only generic rules can be applied and only general statements about the component and turbine condition are possible. Furthermore it showed that the availability and the quality of SCADA signals is crucial when setting up the NBM. The more different SCADA signals are available the higher is the chance of finding correlated signals that can be used as input to predict a certain target signal.

Finally it is important to note that the developed CMS is not aimed to replace vibration monitoring systems. In fact it should rather be seen as an additional tool to gain information about the turbine condition. The great advantage of the system is that no additional sensors are required and that also auxiliary equipment and small components can be monitored provided representative signals exist. In a next project step a vibration analysis tool will be developed and integrated in the CMS. In addition the fleet of the field test will be extended with turbines of a different type and brand.

References

1. Zaher, A., McArthur, S.D.J. and Infield, D.G. (2009). "Online Wind Turbine Fault Detection through Automated SCADA Data Analysis. "*Wind Energy.* 12 pp.574-593.
2. Walford, C. A. (2006). "Wind turbine reliability: understanding and minimizing wind turbine operation and maintenance costs. "*Sandia National Laboraties.* Report no. SAND2006-1100.
3. Hyers, R. W., McGowan, J. G., Sullivan, K. L., Manwell, J. F., Syrett, B. C. (2006). "Condition monitoring and prognosis of utility scale wind turbines. "*Energy Materials Vol.1 No.3.* pp.187-203.
4. Editorial. (2005). "Managing the wind: reducing kilowatt-hour costs with condition monitoring. "*Refocus Vol.6 No.3.* pp.48-51.
5. Kusiak, A. and Verma, A. (2011). "A data-driven approach for monitoring blade pitch faults in wind turbines . "*IEEE Transactions on Sustainable Energly Vol.2 No.1.* pp.87-96.
6. Garlick, W. G., Dixon, R. and Watson, S. J. (2009). "A model-based approach to wind turbine

condition monitoring using SCADA data. "*20th International Conference on Systems Engineering.* 8-10 September, Coventry University, Coventry, UK.
7. Achiche, S., Balazinski, M., Baron, L., Jemielniak, K. (2002). "Tool wear monitoring using genetically-generated fuzzy knowledge bases. "*Engineering Applications of Artificial Intelligence Vol.15.* pp.303-314.
8. Sanz-Bobi, M. A., del Pico, J. and Garcia, M. C. (2006). "SIMAP: Intelligent System for Predictive Maintenance Application to the health condition monitoring of a windturbine gearbox. "Computers in Industry 57 pp.552-568.
9. Zaher, A.S. and McArthur, S.D.J. (2007). "A Multi-Agent Fault Detection System for Wind Turbine Defect Recognition and Diagnosis. "*Proc. IEEE Lausanne POWERTECH.* pp.22-27.
10. Schlechtingen, M. and Santos, I.F. (2011). "Comparative analysis of neural network and regression based condition monitoring approaches for wind turbine fault detection. "*Mechanical Systems and Signal processing Vol.25 No.5.* pp.1849-1875.